Gli Appunti di Archivismi: creare oggetti e correggere errori

(614)—Anche archiviare conoscenza per il futuro può dare problemi; gli Appunti di Archivismi vi possono far risparmiare tempo e karma.

Appunti di Archivismi: creare oggetti e correggere errori



Figure 1:

(614)—Anche archiviare conoscenza per il futuro può dare problemi; gli Appunti di Archivismi vi possono far risparmiare tempo e karma.

30 marzo 2025—Non sono solo rose e fiori quando si lavora su Archive.org. E' un ambiente software particolare, complesso, e gestito da poche persone, quasi tutti volontari.

Per questo motivo, per utilizzarlo aldilà della banale memorizzazione di un singolo file od una singola pagina, è bene avere chiari i fondamentali.

Per chi vuole sapere tutto e subito, consiglio una visita alla pagina della documentazione per gli "sviluppatori", che è molto ben strutturata; non ha una collocazione indipendente, ma si trova nel Blog di Internet Archive.

Nel blog postano coloro che lavorano per Internet Archive, e vi si trovano piccoli gioielli di utilità e di sintesi, tipo questo post dove è descritta la struttura degli oggetti memorizzati su Internet Archive.

Ma per sintetizzare al massimo, Cassandra vi regalerà il distillato dai suoi errori e dai rimedi che ha trovato.

Prima di tutto una raccomandazione; **registratevi subito come utenti.**

E’ ovviamente utile per avere più facilmente accesso agli amministratori, per chiedere più banda per gli inserimenti, etc.

Ma sopratutto, se creaste un oggetto in forma anonima, cioè senza esservi collegato come utente registrato, non lo potreste più modificare, visto che per fare questo occorre essere l’utente creatore dello stesso.

E veniamo ai fondamentali. Internet Archive può essere visto come un enorme singolo database. Gli oggetti che contiene possiedono dei metadati, e tra questi tre sono particolarmente importanti.

L’ identificatore—tutti gli oggetti di Internet Archive sono simili, ovvero sono dello stesso tipo. Ciascun oggetto possiede un’identificatore (un nome univoco) ed almeno un file. Una volta creato, l’oggetto non può essere cancellato, né può esserne modificato l’identificatore.

Se durante la creazione non viene specificato l’identificatore, esso è generato a partire dal nome del primo file caricato. Dopo la creazione, a qualsiasi oggetto possono essere aggiunti altri file od altri metadati.

Non è quindi possibile “recuperare” un identificatore che è stato utilizzato per creare un oggetto, cancellandolo come si farebbe con un semplice file. Infatti, se necessario, un oggetto può essere reso invisibile, ma non può essere mai cancellato.

Morale: un identificatore utilizzato “male” è perso per sempre. Sappiatelo!

Il mediatype—Il mediatype di un oggetto creato dall'interfaccia web viene normalmente dedotto dal tipo del primo file caricato; se l’oggetto invece viene creato da linea comandi usando l’interfaccia IA (che sta per Internet Archive, non altro), è possibile definire il mediatype a piacere, scegliendolo tra i vari tipi previsti.

Se caricate un file che non viene riconosciuto come tipo, al relativo mediatype verrà assegnato il valore più generico possibile, “*data*”, ed il file non potrà essere visualizzato nei visualizzatori standard di Internet Archive.

Il *mediatype*, una volta definito, non può essere più modificato dal proprietario. Se, eccezionalmente, ci sono validi motivi per farlo (tipo, ne avete creati 1000, tutti sbagliati—a Cassandra è capitato!), si può scrivere una email all'amministratore per chiedergli di cambiarlo.

Si deve ovviamente allegare una lista degli identificatori degli oggetti ai quali cambiare il mediatype, oppure una espressione di ricerca che in tutta Internet Archive selezioni solo quegli oggetti.

Se avete scritto con sufficiente gentilezza e precisione, e magari siete già noti per aver versato in precedenza qualche obolo, in un tempo variabile da qualche ora a qualche settimana il vostro desiderio verrà esaudito. Gli archivisti non hanno mai fretta, visto che lavorano per l’eternità.

La collezione (collection)—Non è completamente vero che in Internet Archive tutti gli oggetti sono dello stesso tipo; oltre agli oggetti normali, ne esiste un secondo tipo, detto “C*ollezione*” (collection).

Un oggetto Collezione è semplicemente un “contenitore” di altri oggetti, che non può essere creato dagli utenti normali. Viene usato, ad esempio, per riunire assieme tutti i numeri di una rivista, che siano già stati archiviati come oggetti separati.

Una nuova *collection* deve obbligatoriamente essere richiesta agli amministratori, sempre via email; la richiesta deve provenire da un utente registrato, che abbia già creato un buon numero di documenti. Se i documenti sono già stati creati prima della richiesta di creazione della

collection, è necessario richiedere anche l'assegnazione degli oggetti preesistenti alla nuova collection. Altrimenti dovrete farlo voi, assegnando agli oggetti la nuova collezione come collezione aggiuntiva.

Al momento della creazione, infatti, ad ogni oggetto viene assegnata una collezione primaria; se non ne viene assegnata una esplicitamente, in base al suo mediatype l'oggetto viene inserito nelle collection "community texts" se è un documento, "community videos" se è un video, e così via.

Una volta assegnata, la collezione primaria non può più essere modificata dall'utente. L'utente può eccezionalmente chiederne, come al solito e con i soliti mezzi, una modifica all'amministratore. Tuttavia ce ne è raramente bisogno, perché l'utente può assegnare collection aggiuntive agli oggetti che ha creato. Un oggetto può infatti appartenere a più collezioni.

Qui, ad esempio, trovate tutte le collection di Cassandra, mentre qui trovate tutti gli oggetti da lei creati.

Infine Cassandra vi rinnova una raccomandazione. Non inquinate Internet Archive con le vostre prove. All'atto della creazione è possibile attivare un metadato che crea un oggetto *effimero*, che viene poi cancellato automaticamente e completamente dopo 30 giorni.

Avete il dovere di fare i bravi, quindi usate questo parametro per tutte le archiviazioni di prova, in modo che solo quelle definitive vengano memorizzate permanentemente. Potete ora consultare di nuovo l'intero processo di creazione di un oggetto rileggendo questo articolo, oppure ripassandovi l'intera serie.

A Cassandra non resta che augurarvi "buone archiviazioni"!

Scrivere a Cassandra—Twitter—Mastodon
Videorubrica "Quattro chiacchiere con Cassandra"
Lo Slog (Static Blog) di Cassandra
L'archivio di Cassandra: scuola, formazione e pensiero

***Licenza d'utilizzo**: i contenuti di questo articolo, dove non diversamente indicato, sono sotto licenza Creative Commons Attribuzione—Condividi allo stesso modo 4.0 Internazionale (CC BY-SA 4.0),* tutte le informazioni di utilizzo del materiale sono disponibili a questo link.

By Marco A. L. Calamari on March 31, 2025.

Canonical link

Exported from Medium on August 27, 2025.